ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25dec068

MediaPipe Based Workout Monitoring System Using BlazePose Models

Rajashekar K. J.[1]; Hanumanthappa S.[2]; Chandan S. L.[3]; Ranjitha H.[4]; Ruchitha T. R.[5]; Sinchana S.[6]

[1,2](Professor); [3,4,5,6](Student)

[1,2,3,4,5,6]Department of ISE, Kalpataru Institute of Technology, Tiptur, Karnataka, India

Publication Date: 2025/12/09

Abstract: Real-time exercise analysis can be successfully supported by MediaPipe BlazePose, spatiotemporal models, and machine-learning-driven form classification, according to recent research on pose-estimation-based fitness systems. Existing research shows that lightweight models have potential for posture correction and repetition counting, but it also highlights issues with accuracy, generalization, and user-specific adaptability. In this paper, we offer a MediaPipe-based workout monitoring system that employs the BlazePose model to calculate joint angles, extract 33 body landmarks, and assess exercise form using machine-learning classification and heuristic thresholds. Through an interactive interface, the suggested system offers real-time rep counting, posture feedback, and calorie prediction. With this method, customers can work out precisely and safely without the need for a personal trainer.

Keywords: *KNN, OpenCV, MediaPipoe, Blazepose.*

How to Cite: Rajashekar K. J.; Hanumanthappa S.; Chandan S. L.; Ranjitha H.; Ruchitha T. R.; Sinchana S. (2025) MediaPipe Based Workout Monitoring System Using BlazePose Models. *International Journal of Innovative Science and Research Technology*, 10(12), 129-136. https://doi.org/10.38124/ijisrt/25dec068

I. INTRODUCTION

Nowadays, leading a healthy lifestyle is regarded as one of the most crucial facets of daily life. Although everyone wants to be active and healthy, very few people follow the schedule. Many people are willing to work out, according to our observations, but sometimes they don't know enough about the forms, variants, etc. of the exercises. There are hundreds of different training videos on the internet that offer brief workouts for a variety of hobbies. They are unable to monitor the user's exercise even with these features. It is often observed that even regular gym goers struggle to perform each exercise step perfectly. An action may potentially cause a major long-term harm if it is routinely done wrong. Everyone should engage in moderate-intensity exercise for at least 30 to 60 minutes per day, even those with chronic health conditions. Exercise if they can. Exercise may reduce the risk of acute respiratory distress syndrome, a major cause of death for COVID-19 patients, according to new research. In the same way that COVID-19 has altered healthcare, it has also completely revolutionized how people view physical activity. Gyms and exercise programs can spread illnesses despite being packed spaces with lots of surface area. For this application, we propose to use Mediapipe's flame posture machine learning model, which can operate on a small number of processing-powered devices. The rapid development of deep convolutional neural networks has greatly improved human pose estimation, which helps to accurately analyze the exercise the user is performing. Human pose estimate locates key body locations to accurately identify people's postures in images or movies. This is a necessary prerequisite for the activity.

II. LITERATURE SURVEY

"Human pose estimation" is a technique that may be used in both 2D and 3D to identify stances in photographs or movies. We uncovered numerous algorithms from earlier years that are used for human pose identification. social groups. In order to acquire state-of-the-art results, several of the researchers that developed the new method collaborated with past technology. These days, deep convolutional neural networks deliver leading responses. There are two popular methods: once key-point positions are regressed and key-point heat maps are estimated, the places with the highest heat values are chosen as the key-points.

In order to regress 3D pose from 2D measurements, Ching-Hang Chen's 2017 study first used 2D pose estimate and subsequently 3D exemplar matching using Bigdata sets of 3D mocap data. [1]

Juliet Martinez et al. used a state-of-the-art 2d detector along with a 2d-to-3d human pose estimation in 2017. [2]

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25dec068

A deep-learning multipurpose model for counting repetitions and detecting exercises has been proposed by Abdulmotaleb El Saddik, Fedwa Lamartine, Qingtian Yu, and Haopeng Wang. With 95.69% accuracy in exercise detection on the Rep-Penn dataset, the multitask system can estimate human position, classify physical activities, and count repetitive motions. The multitask model also performed well in repetitive counting, with 0.004 Mean Average Error (MAE) and 0.997 Off-By-One (OBO) accuracy on the Rep-Penn dataset. [3]

In a 2021 study, Choi presented Mobile Human stance, a mobile-friendly model for real-time 3D human stance evaluation from a single RGB image. It comprises of a parametric activation function, a modified MobileNetV2 backbone, and a U-Net-inspired skip concatenation. Compared to the ResNet-50-based model, this model is seven times smaller. Additionally, their extra compact model speeds up inference by 12.2 ms utilizing a Galaxy S20 CPU, which makes it appropriate for real-time 3D human posture assessment in mobile applications. [4]

In 2019, Xie et al. presented the Open Pose human pose estimation algorithm. But its efficiency is relatively low. Deep learning approaches rely on Tensor Flow to identify human body postures in order to get around this. [5]

A paper (by Kumar D. et al.) proposes an Android application for yoga pose estimation that uses the OpenPose model for key landmark detection. The suggested pipeline takes use of a dataset of yoga poses that have been preprocessed using OpenPose and saved on the local system. The pose is then determined by comparing the data from the local machine with the position's real-time recorded event [6].

Unlike typical tests, G. Taware et al. used occlusion-simulating augmentation to test the posture detector in strong occlusions. The estimator finds the user's 33 essential locations prior to employing user alignment. Heatmaps and regression are employed. The test model's layers are trimmed using the techniques from the training model. [7].

A approach for posture estimation leveraging the pre-trained OpenPose model, a multistage CNN model, is provided in a paper by Kanase et al. Important details in videos are recognized by this model. The authors advocate machine learning and a heuristic approach for pose estimation. The heuristic approach makes use of the landmark angles from the OpenPose model. The model uses dynamic time warping to categorize videos. The Euclidean distance between significant locations is calculated and shown to produce a distance matrix. The distance is computed and the best match sites are selected. [8].

III. METHODOLOGY

- *Estimating Human Pose:*

Pose estimation is a machine learning job that uses the spatial locations of particular body components (key points) to predict a person's pose from an image or video. A computer vision method for tracking an object's or person's movements is called pose estimation. This is typically accomplished by locating the items' critical points.

- *Pose Estimation with Deep Learning:*

Deep learning has been demonstrated to perform better than traditional computer vision techniques in a number of tasks, such as object detection and image segmentation, thanks to the quick growth of deep learning solutions in recent years. As a result, posture estimation saw notable improvements and performance increases thanks to deep learning approaches.

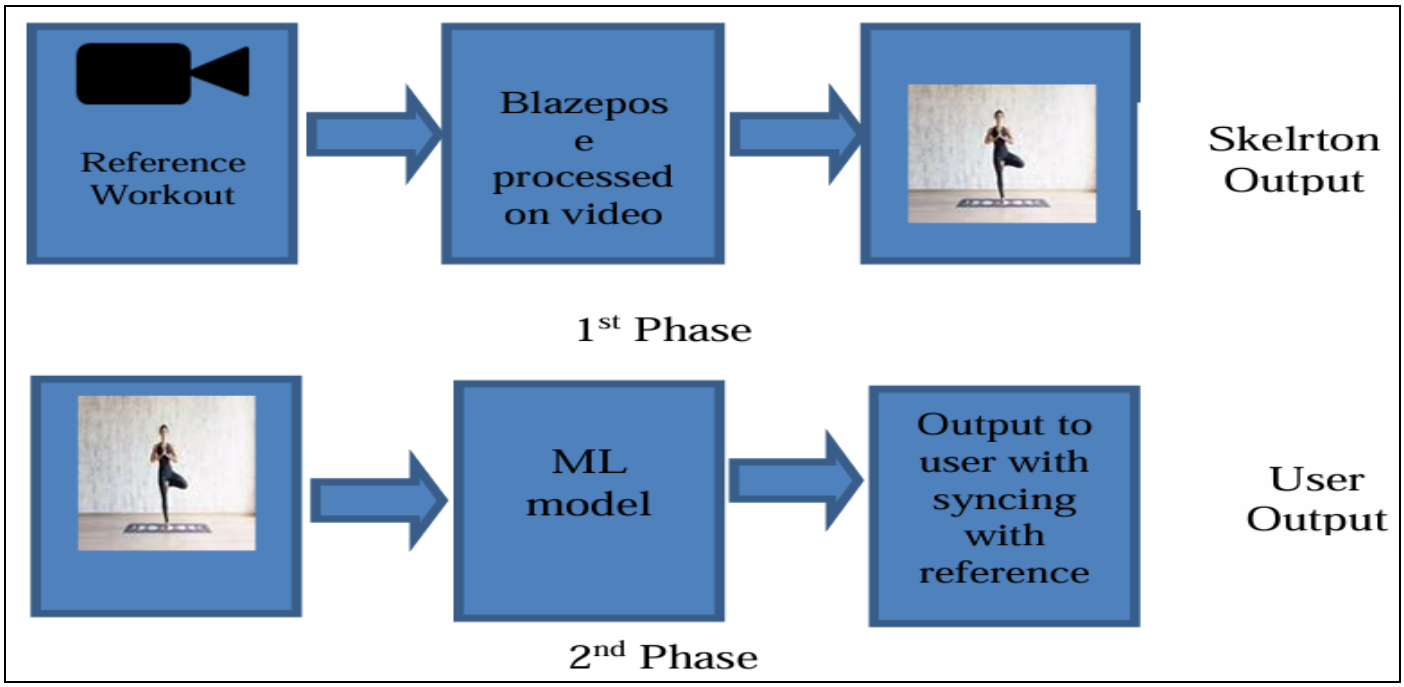

Fig 1 System Architecture

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25dec068

➢ *Stage of Landmark Detection:*

After the subject aligns themself within the camera's field of vision, MediaPipe uses its Pose Landmark Model to identify the region of interest and identify the body landmarks.

➢ *Rep Counting Stage:*

The system uses exercise-specific threshold heuristics and calculates joint angles for every frame. The repetition count is increased if the angles meet the threshold; if not, it stays the same.

➢ *Pose Detection and Angle Calculation:*

Pose detection facilitates monitoring posture throughout any exercise regimen. A collection of predetermined landmarks in pose detection models make it simple to recognize different body parts.

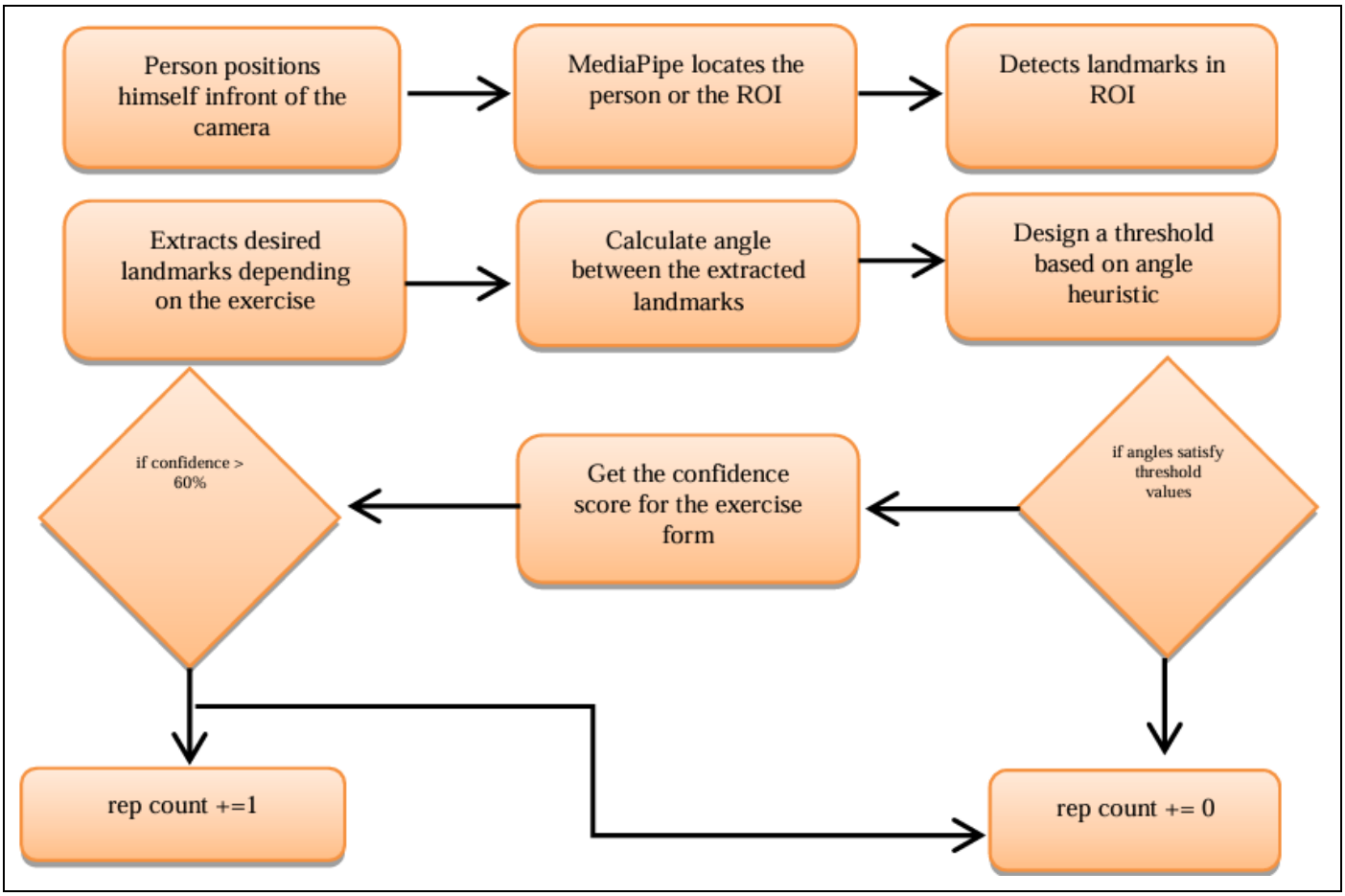

Fig 2 Process Pipeline

The pipeline of the suggested posture estimate, rep counting, and angle calculation model based on heuristic values which are further detailed in depth in the paper—is shown in Fig. 2.

➢ *The BlazePose Model:*

Mediapipe's pose detection API uses BlazePose. The BlazePose variation is suitable for yoga and fitness positions, while performing somewhat poorer than OpenPose. The Pose Detection API is used in the proposed method to identify landmarks. The BlazePose landmark coordinates are shown in Fig. 2. A human position estimate algorithm developed by Google Research can recognize 33 key body parts, including the head, chest, arms, and legs. Because it is fast and accurate, it is ideal for gaming, virtual try-ons, and real-time fitness tracking. BlazePose uses a lightweight neural network architecture with convolutional and depthwise separable convolutional layers for high accuracy and low computing complexity.

In order to determine angles for various workouts, we will extract landmarks, estimate correctness, and count repetitions. Mediapipe's pose detection API is used to implement the BlazePose model. The BlazePose model is a fantastic choice for yoga and fitness poses, although performing somewhat worse than the OpenPose model. The proposed system uses the Pose Detection API to recognize landmarks. Fig. 3 below displays the landmark coordinates from BlazePose's pose and landmark detector model, which provides human posture tracking. A human position estimate algorithm developed by Google Research can recognize 33 key body parts, including the head, chest, arms, and legs.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25dec068

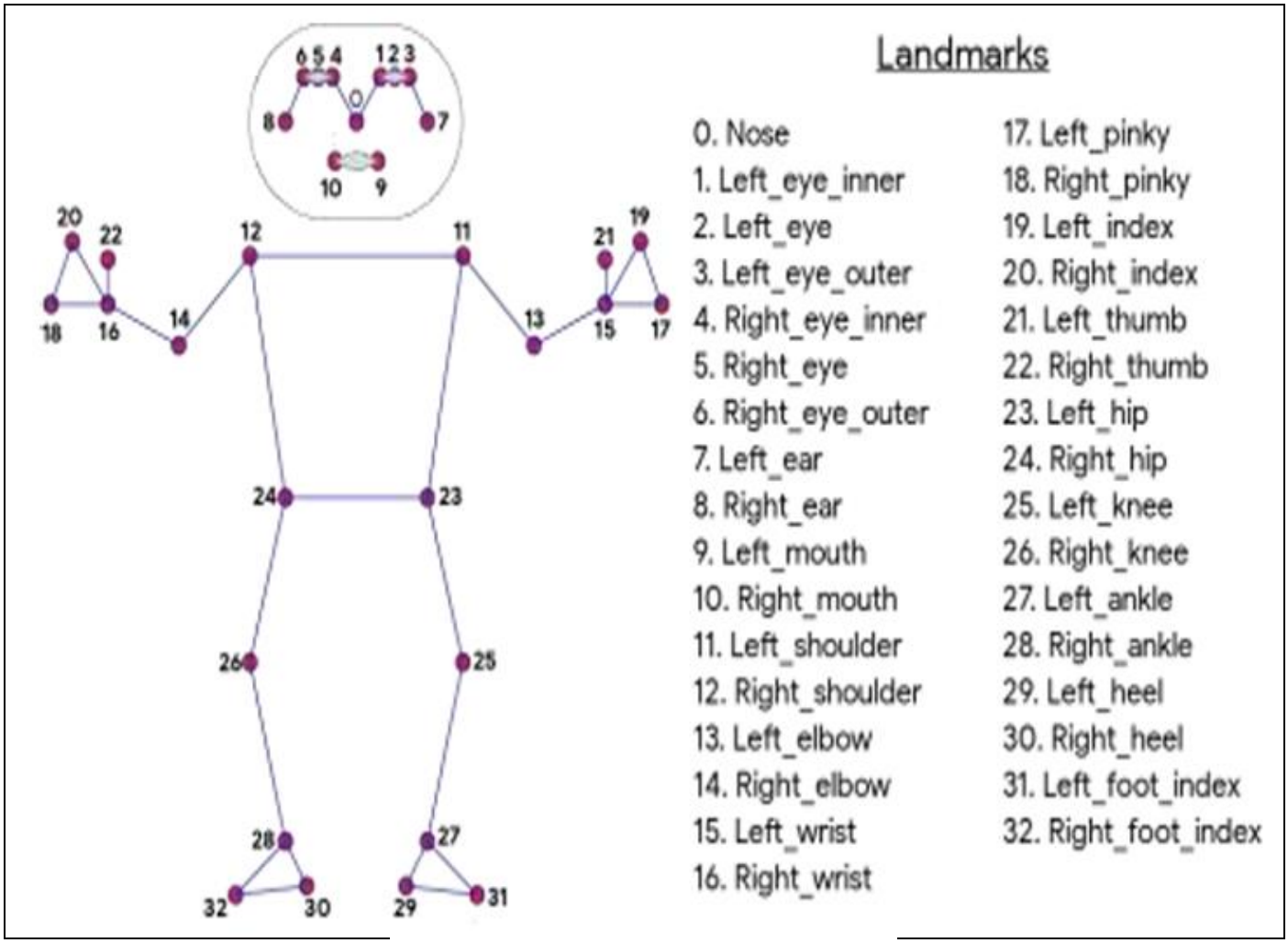

Fig 3 Landmarks of Mediapipe's Blaze Pose Model

➢ *Squat Detection Algorithm:*

It is possible to perform squat analysis by dividing motion into three different domains. upper body, lower body, and movement mechanics. The lower body evaluates hip, knee, and ankle landmark alignment, the upper body focuses on head, neck, and stomach core stability and posture, and movement mechanics evaluates workout length and coordination.

Posture and attitude are computed using the spatial angles of the shoulder, knee, hip, ankle, and foot joints. The back and thigh markers are used to compute the hip angle. The landmarks on the lower leg and thigh dictate the knee angle. The landmark coordinates are extracted using BlazePose's pose and landmark detector model, which provides human posture monitoring of 33 landmarks. In Fig. 4 below, the squat exercise is carried out while MediaPipe tracks the target landmarks in real time.


Fig 4 Squats Exercise Implementation

Table 1 Table of Threshold Heuristic Angle Values for Squats Exercise

Angles	Heuristic Values	Remarks
Hip Angle	50-71 degrees	An imbalance will result from temperatures below 44 degrees.
Knee Angle	55-68 degrees	More than 75 degrees will lessen the impact of the rep.
Torso Angle	35-43 degrees	More over 45 will result in poor posture.
Ankle Angle	75-85 degree	Ankle strain will be excessive if it is less than 75.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25dec068

➢ *The Method for Identifying Bicep Curls:*

Two heuristics are crucial when estimating bicep angle pose. When the angle between the upper arm and body is greater than thirty degrees, the curl is not right. Lifting heavy weights that the body cannot support causes excessive rotation of the upper arms. Initially, the angle between the forearm and upper arm would be between 150 and 180 degrees. If the angle between the upper arm and the forearm is greater than 70 degrees during contraction, the curl is erroneous. Therefore, we use these measurements and threshold values to count the number of repetitions and predict the accuracy accordingly. Thus, we use the shoulder, hip, elbow, and wrist joints to determine the bicep curl stance. Fig. 5 above displays the closed and open angles of the arms during a bicep curl repeat with the landmarks captured by MediaPipe.

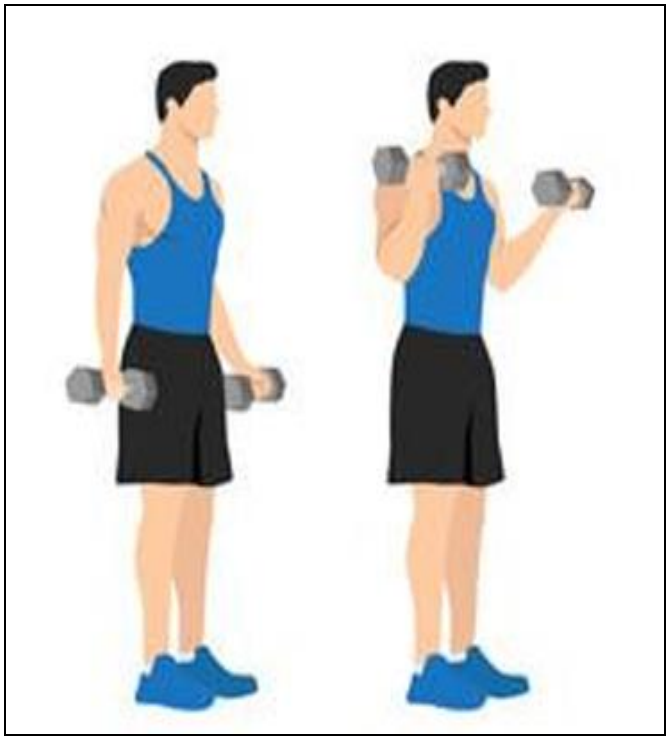

Fig 5 Implementation of the Bicep Curl Exercise (Closed Arms and Open Arms Respectively)

Table 2 Table of Threshold Heuristic Angle Values for Bicep Curls Exercise

Angles	Heuristic Values	Remarks
formed by the upper arm and torso	Not more than 35 degrees	If the angle is greater than 45, the rep will be inaccurate.
formed by the forearm and upper arm when the biceps contract.	Under 70 degrees	More than 70 degrees indicates an unfinished rep.

➢ *What is BlazePose?*

A real-time method for identifying human positions in images or videos is called "blaze pose." It detects a single human position in a single mode. To put it simply, the flame posture is a deep learning model that enables you to estimate human pose by identifying body parts like elbows, hips, wrists, knees, and ankles. It then joins these points to create the skeleton structure of your pose.This lightweight model reduces parameters, computing costs, and improves accuracy by deepening the network via depthwise separable convolution. From the nose to the left foot index, Blaze stance provides us with a total of 33 useful points

➢ *Media Pipe:*

A cross-platform, open-source framework for creating multimodal machine learning pipelines is called Media Pipe. Modern approaches such as multi-hand tracking, object detection, hair segmentation, human face detection, and tracking, and so forth.

IV. RESULTS

When you open the system, you first enter your name, age, and weight. The system then calculates your BMI and recommends suitable exercises based on whether you are underweight, normal, overweight, or obese. Users can watch a demo video to learn the correct exercise form, or they can start the live workout option. When live mode is selected, the camera opens and shows your body with an exoskeleton overlay. As you exercise, the system counts repetitions, shows calories burned, and gives real-time feedback on your posture. If you finish, you can stop the workout by pressing the end button or ‘q’. The system will close the camera and display your accuracy graph and calories burned.If the pose-detection model is not trained, it shows a “not trained” message. The system checks body angles (like hip, knee, feet) during exercises such as squats, compares them with ideal angles, and shows the error in real time. It also calculates calories burned per repetition.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25dec068

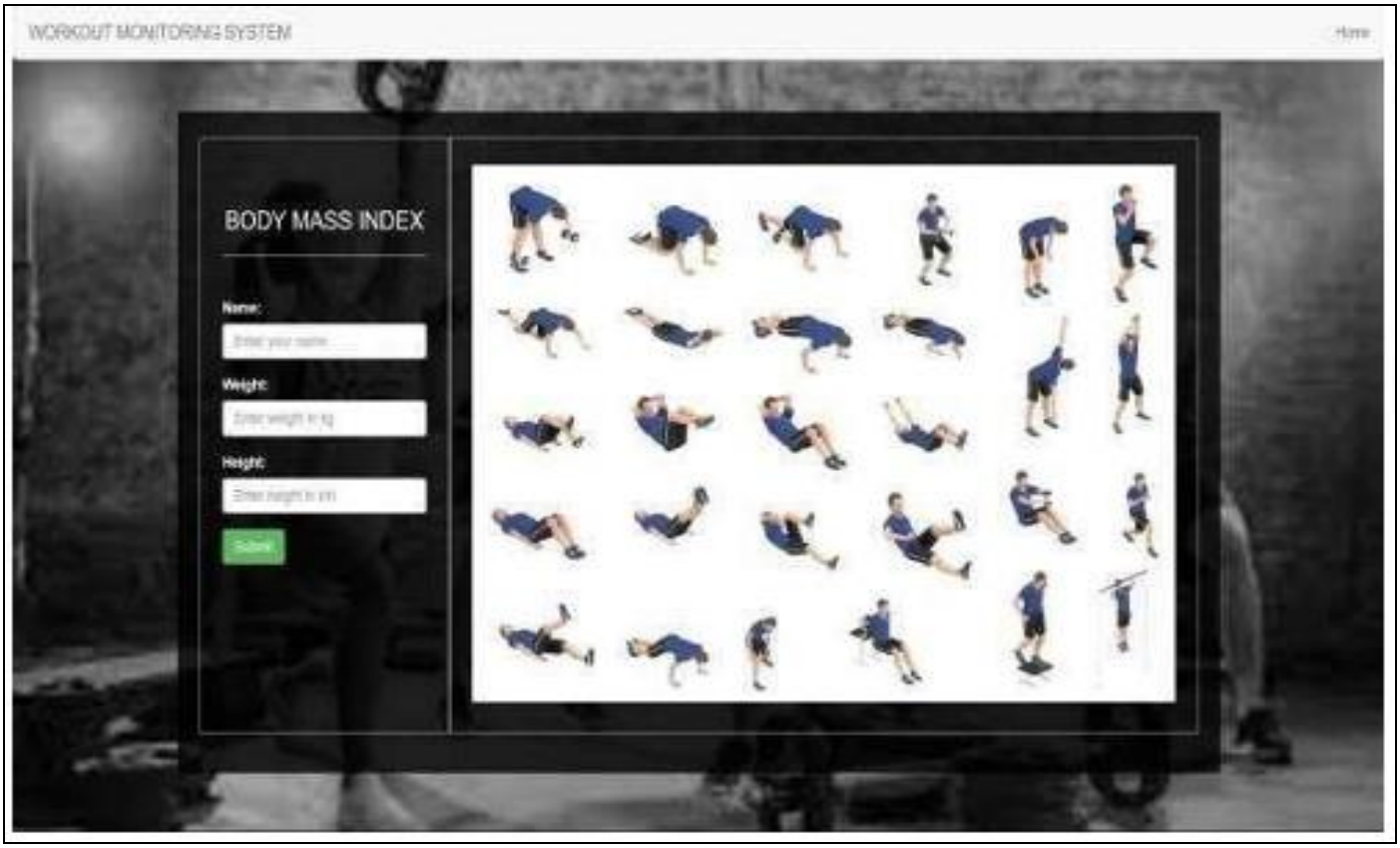

Fig 6 Home Page1

Fig 7 Home Page2

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25dec068

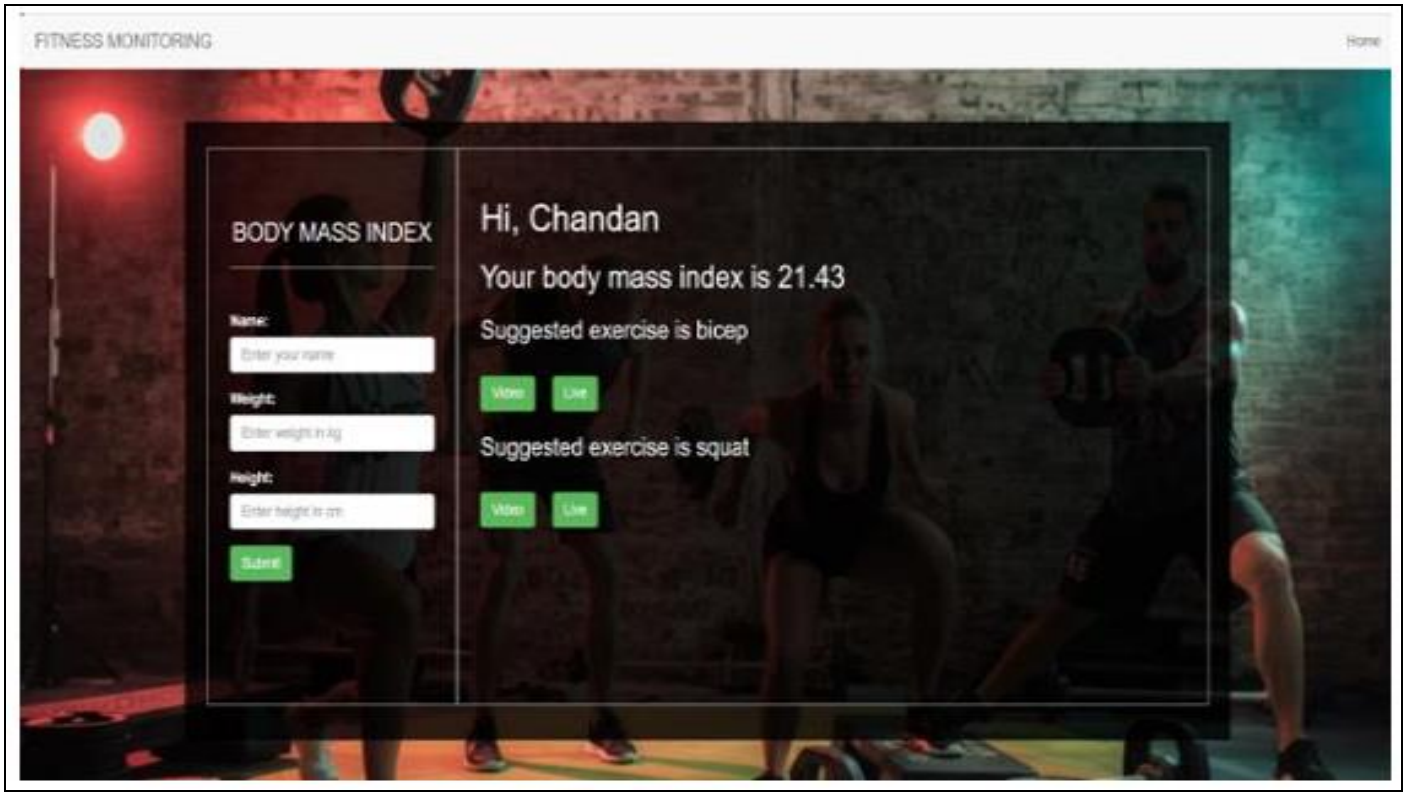

Fig 8 BMI Calculation Page

Fig 9 Live Video Page

V. CONCLUSION

Pose detection has a variety of practical uses. To understand more about pose detection, we examine one such application here. With a high-resolution camera, a tablet or smartphone can do just as much as a desktop or laptop. Pose estimation has advanced to unprecedented levels as mobile phones and tablets have these features. We offer a monitoring application. Exercises without the assistance of a personal trainer. With features like pose prediction, injury avoidance, and real-time workout analysis, the application is producing the greatest outcomes. The system can only be used by one individual at a time for exercise purposes.

REFERENCES

[1]. Ching-Hang Chen and Dev Ramanan 3D Pose Estimation = 2D Pose Estimation + Matching In CVPR,2017

[2]. Martinez, Juliet, et al. "A simple yet effective baseline for 3d human pose estimation." Proceedings of the IEEE International Conference on Computer Vision. 2017.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25dec068

[3]. Qingtian Yu, Haopeng Wang, Fedwa Laamarti and Abdulmotaleb El Saddik, A. Deep Learning-Enabled Multi Task System for Exercise Recognition and Counting. Multimodal Technol. Interact. 2021, 5, 55. https://doi.org/10.3390/mti5090055

[4]. Choi, Sangbum, Seokeon Choi, and Changick Kim. "MobileHumanPose: Toward Real-Time 3D Human Pose Estimation in Mobile Devices." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021. (Choi, 2021)

[5]. Xie, Ling, and Xiao Guo. "Object Detection and Analysis of Human Body Postures Based on TensorFlow." 2019 IEEE International Conference on Smart Internet of Things (SmartIoT). IEEE, 2019. (Xie, 2019)

[6]. G ̈ul, Varol1, Duygu Ceylan2 Bryan Russell2 Jimei Yang BodyNet: Volumetric Inference of 3D Human Body Shapes.

[7]. Zhe cao, Gines Hidalgo, Tomas simon, Shih-En Wei, Yaser Sheikh. OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. Cs.CV, 2018.

[8]. Amit Nagarkoti, Revant Teotia, Amith K. Mahale and Pankaj K. Das. Real time Indoor Workout Analysis Using Machine Learning Computer Vision [2022].